



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06309362 A**(43) Date of publication of application: **04 . 11 . 94**

(51) Int. Cl

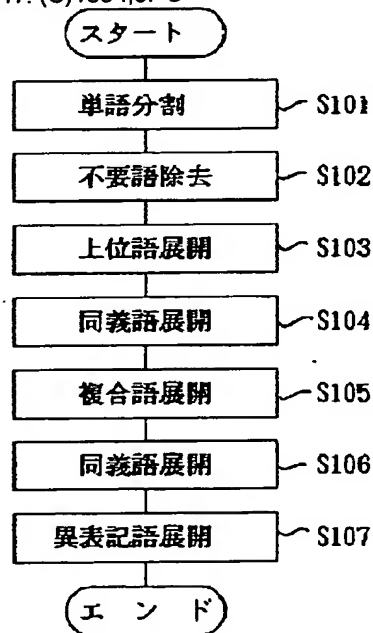
G06F 15/40
G06F 15/40
(21) Application number: **05101348**(22) Date of filing: **27 . 04 . 93**(71) Applicant: **FUJITSU LTD**
(72) Inventor: **KIKUTA YASUYO**
TAZAWA KAZUHIKO
MURATA AKIRA
OTA TATSUO
AKAMATSU YURI
(54) **INFORMATION RETRIEVING METHOD**

COPYRIGHT: (C)1994,JPO

(57) Abstract:

PURPOSE: To shorten the retrieval time while obtaining a high hit rate at the time of retrieval as to the information retrieving means which retrieves information from a book data base for an information processing field by using a key word for retrieval.

CONSTITUTION: When book data are registered, the book data to be registered are divided (101) into words by using a morpheme dictionary, a word which is not necessary as a key word is removed by using an unnecessary word dictionary to employ the result as the key word (102), and a high-order word for the key word is developed (103) by using a high-order word dictionary; and a synonym for the key word which is already generated is developed (104) by using a synonym dictionary, a compound word for the key word which is already generated is divided (105) by using a compound dictionary, and a synonym for the key word generated by dividing the compound word is developed (106) by using the synonym dictionary. Then a new key word is generated (107) by developing a different-notation word for the key word which is already generated by using a different-notation word dictionary.



(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平6-309362

(43)公開日 平成6年(1994)11月4日

(51)Int.Cl.⁵

G 0 6 F 15/40

識別記号

5 0 0 Q 9194-5L

5 1 0 P 9194-5L

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 4 O L (全 13 頁)

(21)出願番号 特願平5-101348

(22)出願日 平成5年(1993)4月27日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中1015番地

(72)発明者 菊田 泰代

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 田澤 和彦

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(72)発明者 村田 彬

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(74)代理人 弁理士 伊東 忠彦

最終頁に続く

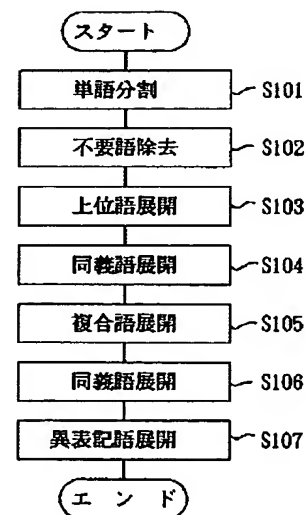
(54)【発明の名称】 情報検索方法

(57)【要約】

【目的】 本発明は、検索用キーワードを用いて情報処理分野の文献データベースより情報を検索する情報検索方法に関し、検索時の高いヒット率を実現しながら、検索時間を短縮することを目的とする。

【構成】 文献データの登録時に、形態素辞書を用いて登録する文献データを単語に分割し(101)、不要語辞書を用いてキーワードとして不要な単語を除去してキーワードとして採用し(102)、上位語辞書を用いて上記キーワードに対する上位語を展開し(103)、同義語辞書を用いて生成済キーワードに対して同義語を展開し(104)、複合語辞書を用いて生成済キーワードに対して複合語を分割し(105)、同義語辞書を用いて複合語の分割で生成されたキーワードに対して同義語を展開し(106)、異表記語辞書を用いて生成済キーワードに対して異表記語を展開して新たなキーワードを生成する(107)。

第1実施例におけるキーワードの展開手順を示すフローチャート



【特許請求の範囲】

【請求項1】 形態素辞書（15）を用いて、文献データベースに登録する文献データの文章を単語に分割するステップ（101）と、

不要語辞書（16）を用いて、上記分割された単語から、検索用キーワードとして不要な単語を除去して、残った単語を検索用キーワードとして採用するステップ（102）と、

上位語辞書（18）を用いて、上記採用されたキーワードに対する上位語を展開して新たな検索用キーワードを生成するステップ（103）と、

同義語辞書（19）を用いて、生成済の検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップ（104）と、

複合語辞書（20）を用いて、生成済の検索用キーワードに対して、複合語を分割して新たな検索用キーワードを生成するステップ（105）と、

上記同義語辞書（19）を用いて、上記複合語の分割により生成された検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップ（106）と、

異表記語辞書（21）を用いて、生成済の検索用キーワードに対して、本来同じ単語であるが表記のみが異なる異表記語を展開して新たな検索用キーワードを生成するステップ（107）とを有し、

上記文献データベースに、上記文献データと共に上記生成した検索用キーワードを登録しておき、文献データの検索時には、検索者により入力された検索用キーワードを用いて、目的とする文献データを検索する構成としたことを特徴とする情報検索方法。

【請求項2】 前記上位語辞書（18）、同義語辞書（19）、複合語辞書（20）、及び異表記語辞書（21）は、検索用キーワードとして有用なものだけが登録されていることを特徴とする請求項1記載の情報検索方法。

【請求項3】 文献データの検索時に、入力された検索用キーワードにより文献データを検索した結果、検索用キーワードに該当する文献データの件数が所定値以上であるかどうかを判断するステップ（203）と、

上記判断の結果、検索用キーワードに該当する文献データの件数が基準値以上である場合には、下位語展開により下位概念の検索用キーワードの候補を生成し、上記検索用キーワードの候補を検索者に提示して、検索者に検索用キーワードを選択させるステップ（204、205）を有することを特徴とする請求項1記載の情報検索方法。

【請求項4】 文献データの検索時に検索者により入力された自然語の質問文を、前記形態素辞書（15）により単語に分割し、前記不要語辞書（16）により不要な

単語を辞去して検索用キーワードとして有用な単語のみを残して、検索用キーワードの候補とするステップ（302、303、304）と、

検索者に、上記検索用キーワードの候補を提示して確認させ、必要に応じて検索用キーワードを変更させて検索用キーワードを決定するステップ（305）とを有し、上記決定された検索用キーワードを用いて文献データを検索することを特徴とする請求項1記載の情報検索方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は情報検索方法に係り、特に、検索用キーワードを用いて情報処理分野の文献データベースより情報を検索する情報検索方法に関する。

【0002】文献データベースより情報を検索する情報検索方法では、目的とする文献データのヒット率の向上、検索時間の短縮が必要とされている。

【0003】

【従来の技術】文献データベースの検索を行う情報検索装置では、文献データの登録時に、登録する文献データにキーワードを付与して情報検索装置に登録しておき、検索時には、キーワードを入力して、目的の文献データを検索している。

【0004】従来、検索時のヒット率を向上させるために、検索時に、同義語辞書、上位語辞書、下位語辞書等の展開語辞書を用いて、検索者により入力された元のキーワードを展開して関連するキーワードを生成し、この生成した関連するキーワードにより検索を行う方法がある。

【0005】

【発明が解決しようとする課題】しかし、従来の、検索時に展開語辞書を用いてキーワードを展開して、生成した関連するキーワードを用いて検索を行う方法では、キーワードの展開に時間がかかるため、検索に要する時間が長いという問題がある。

【0006】また、文献データの登録時には、人手でキーワードを付けるため、ヒット率の向上を可能とする適切なキーワードを付けることが難しく、手間もかかるという問題がある。

【0007】本発明は、上記の点に鑑みてなされたもので、検索時の高いヒット率を実現しながら、検索時間を短縮することができる情報検索方法を提供することを目的とする。

【0008】

【課題を解決するための手段】請求項1の発明は、形態素辞書を用いて、文献データベースに登録する文献データの文章を単語に分割するステップと、不要語辞書を用いて、上記分割された単語から、検索用キーワードとして不要な単語を除去して、残った単語を検索用キーワードとして採用するステップと、上位語辞書を用いて、上

記採用されたキーワードに対する上位語を展開して新たな検索用キーワードを生成するステップと、同義語辞書を用いて、生成済の検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップと、複合語辞書を用いて、生成済の検索用キーワードに対して、複合語を分割して新たな検索用キーワードを生成するステップと、上記同義語辞書を用いて、上記複合語の分割により生成された検索用キーワードに対して、読みは異なるが同じ意味の同義語を展開して新たな検索用キーワードを生成するステップと、異表記語辞書を用いて、生成済の検索用キーワードに対して、本来同じ単語であるが表記のみが異なる異表記語を展開して新たな検索用キーワードを生成するステップとを有し、上記文献データベースに、上記文献データと共に上記生成した検索用キーワードを登録しておき、文献データの検索時には、検索者により入力された検索用キーワードを用いて、目的とする文献データを検索する構成とする。

【0009】請求項3の発明は、文献データの検索時に、入力された検索用キーワードにより文献データを検索した結果、検索用キーワードに該当する文献データの件数が所定値以上であるかどうかを判断するステップと、上記判断の結果、検索用キーワードに該当する文献データの件数が基準値以上である場合には、下位語展開により下位概念の検索用キーワードの候補を生成し、上記検索用キーワードの候補を検索者に提示して、検索者に検索用キーワードを選択させるステップを有する構成とする。

【0010】請求項4の発明は、文献データの検索時に検索者により入力された自然語の質問文を、前記形態素辞書により単語に分割し、前記不要語辞書により不要な単語を除去して検索用キーワードとして有用な単語のみを残して、検索用キーワードの候補とするステップと、検索者に、上記検索用キーワードの候補を提示して確認させ、必要に応じて検索用キーワードを変更させて検索用キーワードを決定するステップとを有し、上記決定された検索用キーワードを用いて文献データを検索する構成とする。

【0011】

【作用】請求項1の発明では、文献データの登録時に、展開語辞書により自動的に適切な検索用キーワードを生成して登録する。このため、検索時の高いヒット率を実現しながら、検索時間を短縮することを可能とし、また、検索用キーワードを付与する手間を削減することを可能とする。

【0012】請求項3の発明では、検索用キーワードによる文献データの検索の結果、該当する文献データの件数が基準値以上の場合には、下位概念の検索用キーワード候補を生成して検索者に選択させて、検索対象を絞り込む。このため、検索用キーワードに該当する文献デー

タの件数が多い場合に、目的の文献データを検索するまでの検索時間を短縮することを可能とする。

【0013】請求項4の発明では、検索時に、検索者により入力された自然語の質問文から、適切なキーワードを生成する。このため、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることを可能とする。

【0014】

【実施例】図1は本発明の第1実施例を適用した情報検索装置の構成図を示す。この情報検索装置は、情報処理分野の文献データベースを登録し、検索する装置の例である。図1に示す情報検索装置は、登録する文献データを入力する文献データ入力装置12、検索者が検索用キーワードやコマンド等を入力する入力装置13、検索結果等を表示する表示装置14、文献データとキーワードの登録処理、検索処理を行う処理部11を備えている。

【0015】また、補助記憶装置等に格納されるデータベース25は、複数の文献データからなる文献データファイル26と、各文献データに付与された検索用キーワードの集まりでありキーワードファイル27からなる。キーワードファイル27中の各キーワードは、対応する文献データの格納位置を示すポインタと共に記憶されている。

【0016】また、文献データ登録時に、文献データの文章からキーワードを抽出して、抽出した各キーワードを基に、関連する単語にキーワードを展開して検索用キーワードを生成するために、形態素辞書15、不要語辞書16、展開語辞書17を備えている。展開語辞書17は、上位語辞書18、同義語辞書19、複合語辞書20、異表記語辞書21からなる。上記の各辞書は、補助記憶装置等に格納されている。

【0017】形態素辞書15には、各種文字列と品詞(名詞、接続詞、助詞等)の対応を示す情報が記憶されている。不要語辞書16には、接続詞、助詞等の不要な単語、及び、情報処理分野でキーワードとして無用な単語が登録されている。例えば、“ソフト”という単語は、極一般的単語であるため、不要語として登録されている。

【0018】また、上位語辞書18には、情報処理分野で使われる各種の単語について、上位概念の単語、下位概念の単語を示す情報が記憶されている。同義語辞書19には、情報処理分野で使われる各種の単語について、読みは異なるが同じ意味の同義語を示す情報が記憶されている。

【0019】複合語辞書20には、情報処理分野で使われる各種複合語について、複合語を分割した単語が登録されている。

【0020】異表記語辞書21には、情報処理分野で使われる各種単語について、本来同じ単語であるが表記のみが異なる異表記語を示す情報が記憶されている。互い

に異表記語の関係にあるものとしては、例えば、一つの単語を、アルファベットで表記した単語と、カタカナで表記した単語がある。

【0021】なお、上記展開語辞書17には、キーワードとして無用なものは、登録していない。

【0022】次に、第1実施例における、文献データを登録する際のキーワードの展開について説明する。本実施例では、文献データを登録する際に、処理部11が、文献データの文章から自動的にキーワードを抽出して、抽出した各キーワードを基に、関連する単語にキーワードを展開して検索用キーワードを生成する。

【0023】図2は、第1実施例におけるキーワードの展開手順を示すフローチャートである。また、図3は、第1実施例におけるキーワードの展開の一例の説明図を示す。

【0024】まず、文献データ入力装置12により入力された文献データの文章を、形態素辞書15を用いて単語に分割する(ステップ101)。次に、不要語辞書16を用いて、接続詞、助詞等の不要な単語、及び、情報処理分野におけるキーワードとして無用な単語を辞去する。この不要語の除去が終了した時点で残った単語を、文献データの文章から抽出されたキーワードとして採用する(ステップ102)。

【0025】この後に続くキーワード展開処理では、ステップ102で抽出された各キーワードに対して、順次、各展開処理を行って、新たなキーワードを生成する。まず、上位語辞書18を用いて、元のキーワードの上位語を展開して新たなキーワードを生成する。図3の例では、上位語展開113に示すように、元のキーワードである、“LOTUS1-2-3”に対して、上位語である“スプレッドシート”を新たなキーワードとして生成している(ステップ103)。

【0026】次に、同義語辞書19を用いて、ステップ103までに生成済のキーワードに対して同義語を展開して、新たなキーワードを生成する。図3の例では、同義語展開114に示すように、“スプレッドシート”に対する同義語である、“表計算ソフト”を新たなキーワードとして生成している(ステップ104)。

【0027】次に、複合語辞書20を用いて、ステップ104までに生成済のキーワードに対して複合語を分割して、新たなキーワードを生成する。図3の例では、複合語展開115に示すように、複合語である“表計算ソフト”に対して、分割した“表計算”と、“計算ソフト”を新たなキーワードとして生成している。なお、“表計算ソフト”を分割すると“表”、“計算”、“ソフト”も生じるが、これらは、情報処理分野のキーワードとして無用であるので、複合語辞書20には、登録していない。このように、複合語を分割して生成される単語のうち、キーワードとして無用なものを登録しないでおくことにより、余分なキーワードを生成することを防

ぐことができ、キーワードファイル27の容量の削減もできる(ステップ105)。

【0028】次に、再度、同義語辞書19を用いて、ステップ105で複合語の分割により生成されたキーワードに対して同義語を展開して、新たなキーワードを生成する。図3の例では、同義語展開116に示すように、“表計算”に対する同義語である、“カルク”を新たなキーワードとして生成している(ステップ106)。

【0029】次に、異表記語辞書21を用いて、ステップ106までに生成済のキーワードに対して異表記語を展開して、新たなキーワードを生成する。図3の例では、異表記語展開117に示すように、“LOTUS1-2-3”に対する異表記語の“ロータス1-2-3”、及び“スプレッドシート”に対する異表記語の“SPREADSHEET”を、新たなキーワードとして生成している(ステップ107)。

【0030】上記の展開辞書17を用いたキーワード展開で生成されたキーワードは、登録する文献データのキーワードとして、キーワードファイル27に登録される。一方、文献データは、データベース25の文献データファイル26に登録される。

【0031】上記のように、本実施例では、文献データの登録時に、展開語辞書17を用いて、文献データから抽出した元のキーワードと関連する有用なキーワードを自動的に生成して登録する。このため、文献データの検索時には、キーワードの展開処理をすることなく、広い範囲のキーワードで目的の文献を検索することができる。

【0032】従って、本実施例では、文献データの検索時に、従来方法に比べて、キーワード展開処理に要する時間分、検索時間を短縮することができ、かつ、良好なヒット率を実現することができる。また、人手により検索用キーワードを付与する手間を削減することができる。

【0033】次に、第1実施例における、文献データの検索について説明する。図4は、第1実施例における検索の手順を示すフローチャートである。

【0034】キーワードが検索者により入力されると(ステップ201)、入力されたキーワードを用いて、データベース25より文献データの検索を行う(ステップ202)。次に、検索の結果、キーワードに該当する文献データの件数が基準値以上であるかどうかを判断する(ステップ203)。

【0035】検索結果の件数が基準値以上である場合には、件数を絞るために、下記の処理を行う。まず、上位語辞書18を逆引きして、元のキーワードの下位語を展開して、下位概念のキーワードを生成する(ステップ204)。次に、生成した下位概念のキーワードの一覧を表示装置14により表示して、検索者に使用する下位概念のキーワードを選択させる(ステップ205)。

【0036】次に、ステップ202に戻って、ステップ205で選択された、元のキーワードの下位概念のキーワードを用いて、再度文献データの検索を行う。以後、検索結果の件数が基準値未満になるまで、上記の処理を続ける。

【0037】ステップ203で、検索結果の件数が基準値未満である場合は、検索結果の文献データの一覧を表示装置14に表示して、検索者に文献データを選択させる(ステップ206)。検索者により文献データが選択されると、この選択された文献データを表示装置14等

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636

【0053】上記キーワードの選択肢を表示する時点では、処理部11により、第1候補として、「食品卸」、「物流管理」、「事例」が選択されている。

【0054】また、データベースの選択肢には、上記各検索キーで現在選択されている選択肢に関係するデータベースの種類が表示される。図8(B)では、「事例データベース」、「提案書事例データベース」、「開発事例データベース」～が表示されている。また、第1候補として、処理部11により、「事例データベース」が選択されている。

【0055】検索者は、上記例のように表示されたキーワードの選択肢とデータベースの選択肢を選択して、検索条件を決定する。

【0056】検索者がキーワードの選択肢とデータベースの選択肢を選択して、検索条件が決定されると、処理部11は、文献データの検索を行う。即ち、データベース定義情報22を基に、選択されたデータベースを調べて、決定された検索条件に従って、文献データの検索を行う(ステップ306)。この検索の結果、各検索キーのキーワードの組合せに該当する文献データの件数を表示する(ステップ307)。

【0057】次に、検索者に、検索範囲の拡大又は絞り込みのために、検索条件の変更をするかどうかの確認をさせる。検索の結果、該当件数が少なく、検索者が検索条件の変更を選択し、かつ、検索範囲の拡大を選択した場合には、展開語辞書17を用いてキーワードの上位語を展開して、展開した上位概念のキーワードを検索者に選択させる。

【0058】また、検索の結果、該当件数が多く、検索者が検索条件の変更を選択し、かつ、検索範囲の絞り込みを選択した場合には、図4の検索手順で示したように、展開語辞書17を用いてキーワードの下位語を展開して、展開した下位概念のキーワードを検索者に選択させる。

【0059】検索条件が変更されて、検索者により新たなキーワードが選択された場合には、ステップ306に戻って、再度、文献データの検索を行う。

【0060】ステップ308で、検索者が検索条件の変更を選択しなかった場合には、検索結果の文献データの一覧を表示装置14に表示して、検索者に文献データを選択させる(ステップ310)。検索者により文献データが選択されると、この選択された文献データを表示装置14等に出力する(ステップ311)。

【0061】図8の例では、図8(B)に示す検索条件で検索を行った結果、図8(C)に示すように、検索結果の表示がされている。この場合、該当件数が1件で少ないため、図8(C)に示すように、検索者は、検索条件の変更を選択して、検索範囲の拡大を選択し、範囲を広げたい検索キーとして、業種を選択している。

【0062】これにより、処理部11は、業種の元のキ

ーワードである「食品卸」を上位語に展開して、上位概念のキーワードの一覧である、「流通」、「販売」～を表示している。検索者は、この表示された上位概念のキーワードの一覧から、キーワードを選択する。これにより、検索対象を拡大して、再度、検索を行うことができる。

【0063】また、図9の例では検索者により図9

(A)に示す、自然語の質問文、「UNIXに関するマニュアルとアプリケーションガイドを探したい」が、入力されている。これに対して、図9(B)に示すように、各検索キーごとのキーワードの選択肢と、データベースの選択肢を表示している。

【0064】検索キーのOSには、キーワードの選択肢として、「UNIX」だけが表示され、他の検索キーには、キーワードの選択肢として、「マニュアル」と「アプリケーションガイド」が表示されている。なお、情報検索装置により、第1候補として、「マニュアル」が選択されている。

【0065】また、データベースの選択肢には、「マニュアルデータベース」、「入門書データベース」～が表示されている。また、第1候補として、処理部11により、「マニュアルデータベース」が選択されている。

【0066】検索者は、上記例のように表示されたキーワードの選択肢とデータベースの選択肢を選択して、検索条件を決定する。

【0067】図9の例では、図9(B)に示す検索条件で検索を行った結果、図9(C)に示すように、検索結果の表示がされている。この場合、該当件数が68件と多いため、図9(C)に示すように、検索者は、検索条件の変更を選択して、検索範囲の絞り込みを選択し、絞り込みたい条件として、OSを選択している。

【0068】これにより、処理部11では、OSの元のキーワードである「UNIX」を下位語に展開して、下位概念のキーワードの一覧である、「UTS/M」、「UXP/M」、「SX/A」、「UXP/DS」、「SUN/OS」～を表示している。検索者は、この表示された下位概念のキーワードの一覧から、キーワードを選択する。これにより、検索対象を絞り込んで、再度、検索を行うことができる。

【0069】上記のように、第2実施例では、検索時に、検索者により入力された自然語の質問文から、キーワードの選択肢とデータベースの選択肢を生成し、検索者にこの選択肢を選択させて決定した検索条件で、文献データの検索を行う。このため、検索時に、SQL(Structured Query Language)等のような所定の質問式による記述が不要で、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることができる。

【0070】

【発明の効果】上述の如く、請求項1の発明によれば、

10

20

30

40

50

文献データの登録時に、展開語辞書により自動的に適切な検索用キーワードを生成して登録するため、検索時の高いヒット率を実現しながら、検索時間を短縮することができ、また、検索用キーワードを付与する手間を削減することができる等の特長を有する。

【0071】請求項3の発明によれば、検索用キーワードによる文献データの検索の結果、該当する文献データの件数が基準値以上の場合には、下位概念の検索用キーワード候補を生成して検索者に選択させて、検索対象を絞り込むため、検索用キーワードに該当する文献データの件数が多い場合に、目的の文献データを検索するまでの検索時間を短縮することができる。

【0072】請求項4の発明では、検索時に、検索者により入力された自然語の質問文から、適切なキーワードを生成するため、容易に適切なキーワードを生成させることができ、検索時のヒット率を向上させることができる。

【図面の簡単な説明】

【図1】本発明の第1実施例を適用した情報検索装置の構成図である。

【図2】第1実施例におけるキーワードの展開手順を示すフローチャートである。

【図3】第1実施例におけるキーワード展開の一例の説明図である。

【図4】本発明の第1実施例における、検索手順を示すフローチャートである。

【図5】第1実施例の検索時における下位概念のキーワードの生成例を示す図である。

【図5】

第1実施例の検索時における
下位概念のキーワードの生成例を示す図

	元のキーワード	生成されたキーワード
A	製 造	化 学 機 械 }
B	保 險	生命保険 損害保険 }
C	パソコン	FMRシリーズ FM-TOWNS シリーズ }

* 【図6】本発明の第2実施例を適用した情報検索装置の構成図である。

【図7】本発明の第2実施例における、検索手順を示すフローチャートである。

【図8】自然語の質問文の例と、検索条件確認時及び検索条件変更時の表示例を示す図である。

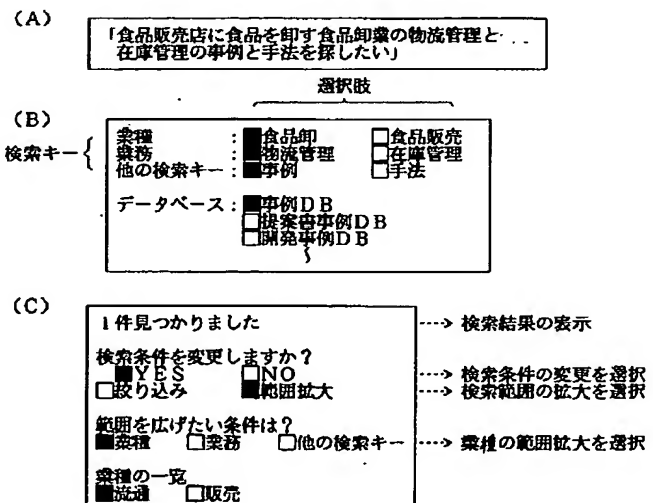
【図9】自然語の質問文の例と、検索条件確認時及び検索条件変更時の表示例を示す図である。

【符号の説明】

- 1 1 処理部
- 1 2 文献データ入力装置
- 1 3 入力装置
- 1 4 表示装置
- 1 5 形態素辞書
- 1 6 不要語辞書
- 1 7 展開語辞書
- 1 8 上位語辞書
- 1 9 同義語辞書
- 2 0 複合語辞書
- 2 1 異表記語辞書
- 2 2 データベース定義情報
- 2 3 知識ベース
- 2 4 データベース群
- 2 5 データベース
- 2 5₁ ~ 2 5₅ データベース
- 2 6 文献データファイル
- 2 7 キーワードファイル

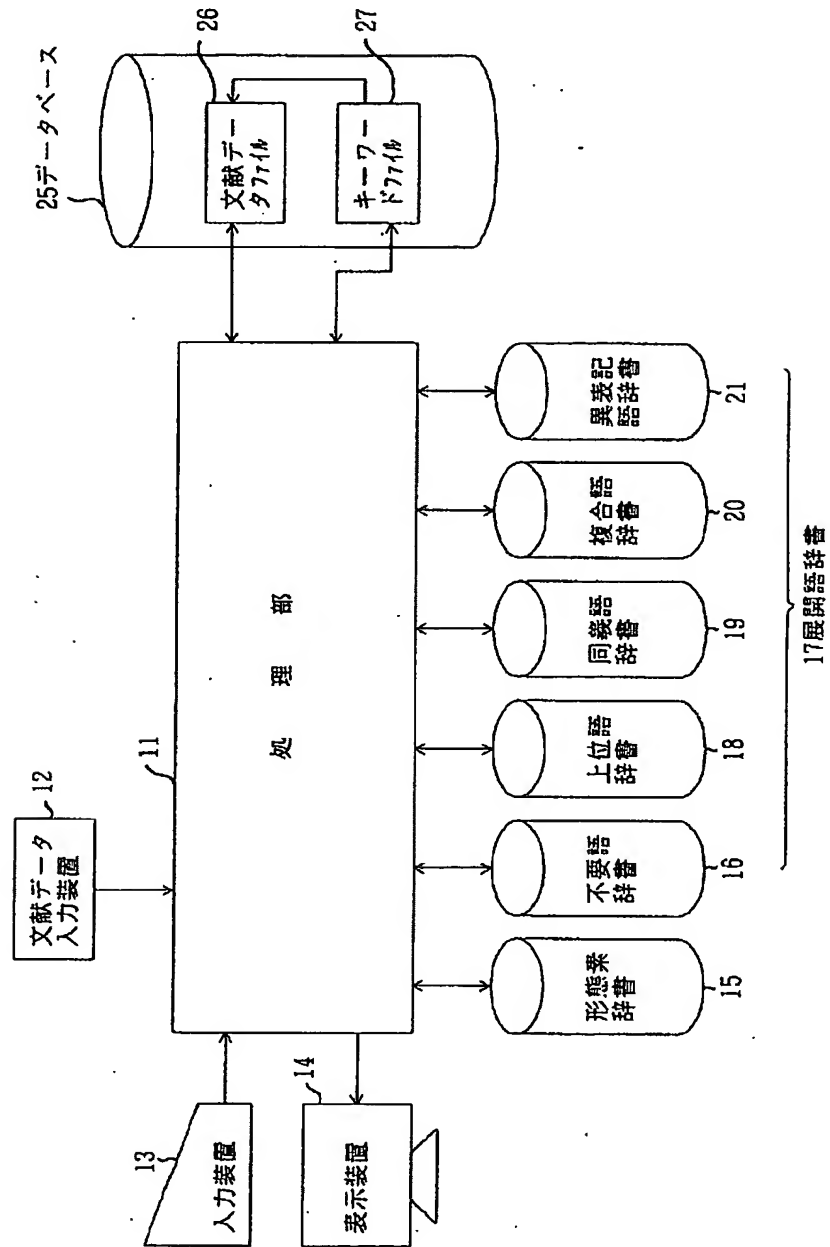
【図8】

自然語の質問文の例と、検索条件確認時及び
検索条件変更時の表示例を示す図



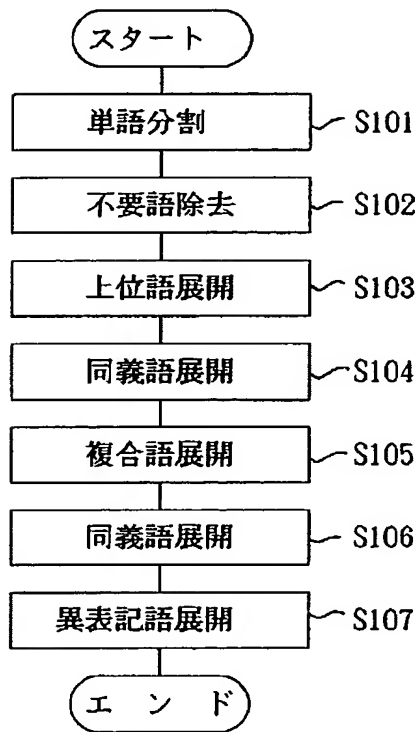
【図1】

本発明の第1実施例を適用した情報検索装置の構成図



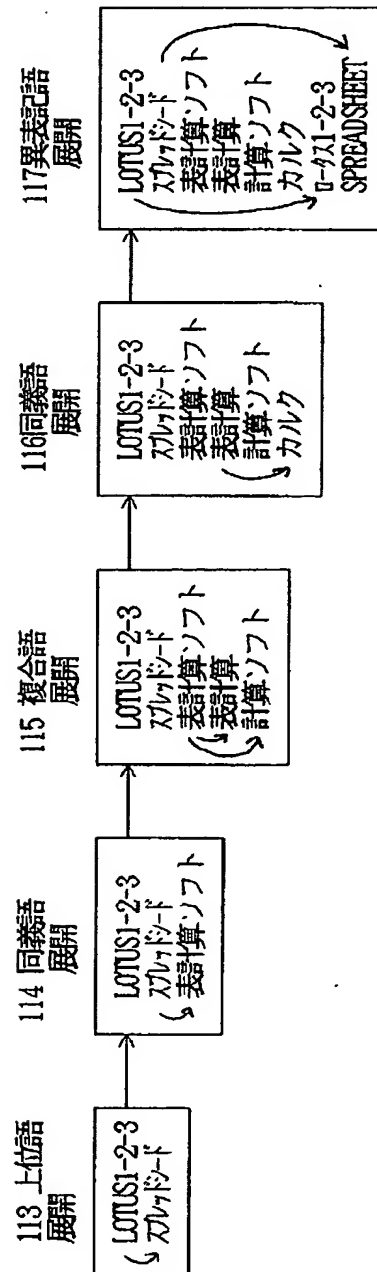
【図2】

第1実施例におけるキーワードの
展開手順を示すフローチャート



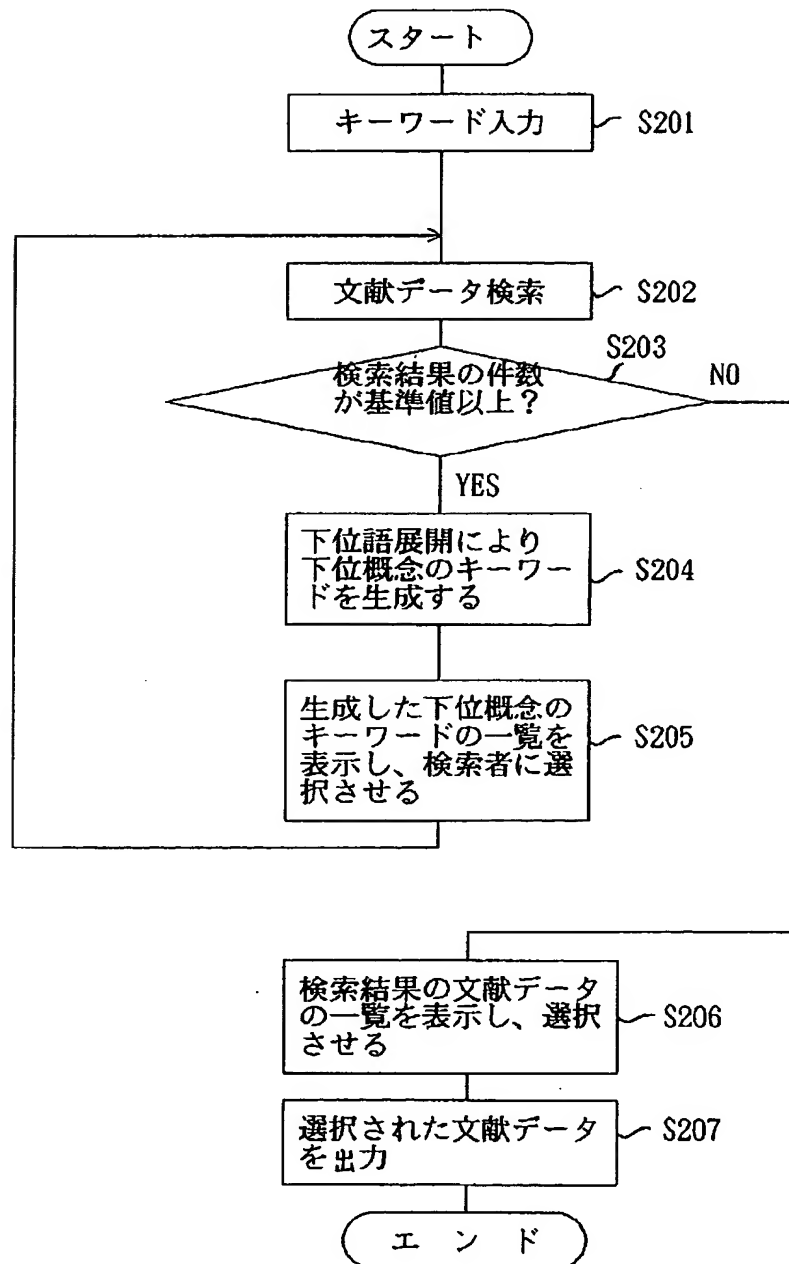
【図3】

第1実施例におけるキーワード展開の一例の説明図



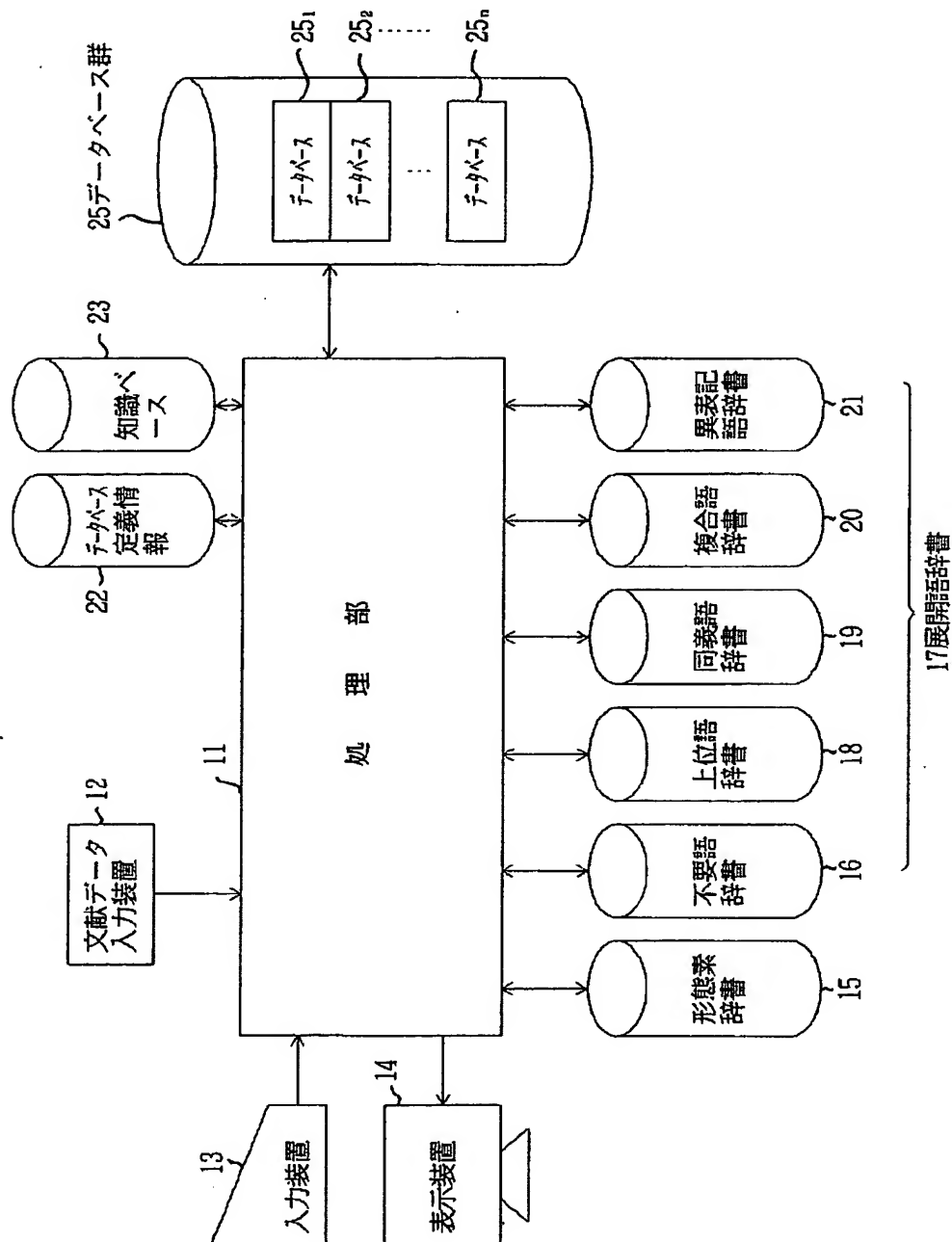
【図4】

本発明の第1実施例における
検索手順を示すフローチャート



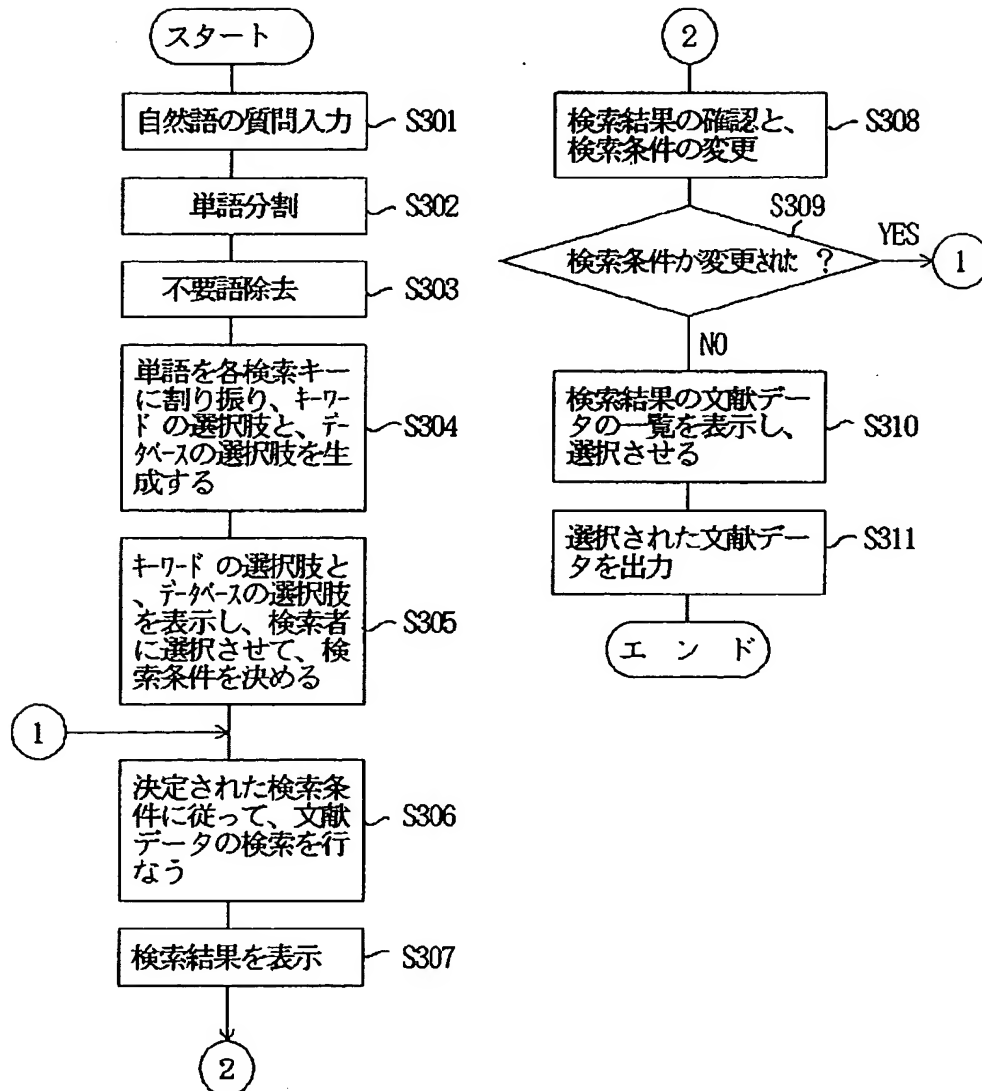
【図6】

本発明の第2実施例を適用した情報検索装置の構成図



【図7】

本発明の第2実施例における
検索の手順を示すフローチャート



【図9】

自然語の質問文の例と、検索条件確認時及び
検索条件変更時の表示例を示す図

(A) 「UNIXに関するマニュアルとアプリケーションガイドを探したい」

選択肢

(B) 検索キー { OS : ☒ UNIX
他の検索キー : ☒ マニュアル ☐ アプリケーションガイド
データベース : ☒ マニュアルDB
 ☐ 入門書DB }

(C) 68件見つかりました ---> 検索結果の表示
検索条件を変更しますか？
☒ YES ☐ NO ---> 検索条件の変更を選択
☐ 絞り込み ☒ 検索範囲 ---> 検索範囲の絞り込みを選択
絞り込みたい条件は？
☒ OS ☐ 他の検索キー ---> OSの絞り込みを選択
OSの一覧
☐ ITS/M ☐ LRP/M ☐ SX/A
☐ LRP/DS ☒ SUN/OS }

フロントページの続き

(72)発明者 太田 竜夫
神奈川県川崎市中原区上小田中1015番地
富士通株式会社内

(72)発明者 赤松 由理
神奈川県川崎市中原区上小田中1015番地
富士通株式会社内